JVI Journals.ASM.org

Structure of the C-Terminal Domain of Lettuce Necrotic Yellows Virus Phosphoprotein

Nicolas Martinez,[a,b,c,d] Euripedes A. Ribeiro, Jr.,[a,b,c] Cédric Leyrat,[e] Nicolas Tarbouriech,[a,b,c] Rob W. H. Ruigrok,[a,b,c] Marc Jamin[a,b,c]

Université Grenoble Alpes, UVHCI, Grenoble, France[a]; CNRS, UVHCI, Grenoble, France[b]; Unit for Virus Host-Cell Interactions, Université Grenoble Alpes-EMBL-CNRS, Grenoble, France[c]; Institut Laue Langevin, Grenoble, France[d]; Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, United Kingdom[e]

Lettuce necrotic yellows virus (LNYV) is a prototype of the plant-adapted cytorhabdoviruses. Through a meta-prediction of disorder, we localized a folded C-terminal domain in the amino acid sequence of its phosphoprotein. This domain consists of an autonomous folding unit that is monomeric in solution. Its structure, solved by X-ray crystallography, reveals a lollipop-shaped structure comprising five helices. The structure is different from that of the corresponding domains of other *Rhabdoviridae*, *Filoviridae*, and *Paramyxovirinae*; only the overall topology of the polypeptide chain seems to be conserved, suggesting that this domain evolved under weak selective pressure and varied in size by the acquisition or loss of functional modules.

Numerous viruses infecting plants are classified as rhabdoviruses most generally on the basis of electron microscopic observations of distinctive enveloped bacilliform or bullet-shaped particles in the infected cells (1, 2). Only a few of them were confirmed to be true rhabdoviruses by sequencing of their genome and molecular studies (1, 2). The *Rhabdoviridae* belong to the nonsegmented negative-stranded RNA viruses (NNVs); together with the *Paramyxoviridae*, the *Filoviridae*, and the *Bornaviridae*, they constitute the order *Mononegavirales*. Plant rhabdoviruses are classified in two genera on the basis of their site of virion maturation; members of the genus *Cytorhabdovirus* assemble in the cytoplasm, whereas members of the genus *Nucleorhabdovirus* assemble in the nucleus (3).

The lettuce necrotic yellows virus (LNYV), identified in 1963 in Australia (4), is found only in Australia and New Zealand, but it is very similar to the lettuce yellow mottle virus that was recently isolated in France (5). Infection by LNYV causes fading of the leaves and necrosis of the heads, rendering the lettuce unsuitable for consumption. LNYV is transmitted by an aphid of the genus *Hyperomyzus* (4), and the main natural reservoir of both virus and vector is the common sowthistle (*Sonchus oleraceus*), which is infected without showing any symptoms. Disease outbreaks occur where the natural reservoir is densely populated and coincide with flight periods of the vector.

LNYV is the prototype member of the *Cytorhabdovirus* genus (3). In the late stages of LNYV replication, viral particles bud from the endoplasmic reticulum and accumulate in the cytoplasm (6, 7). The genome of LNYV was completely sequenced, revealing an organization conserved among other rhabdoviruses (8). It consists of six genes flanked by untranslated 3′ leader and 5′ trailer sequences, and intergenic regions contain highly conserved consensus sequences. On the basis of sequence analysis and immunolabeling, the genes were proposed to encode successively from the 3′ extremity to the 5′ extremity the nucleoprotein (N), the phosphoprotein (P), a plant-specific 4b protein assisting in cell-to-cell movement of viral complexes through plasmodesmata and the vascular system, the matrix protein (M), the glycoprotein (G), and the large subunit of the RNA-dependent RNA polymerase (L) (8–12).

In infected cells, the N and P proteins of LNYV interact with each other, as in plant nucleorhabdoviruses and other rhabdoviruses (6), but little is known about the functions of P in LNYV or in other members of the *Cytorhabdoviridae*. However, the structure and function of the P protein from rhabdoviruses have been studied, in particular, those of the P protein from vesicular stomatitis virus (VSV), a *Vesiculovirus*, and from rabies virus (RAV), a *Lyssavirus* (13). In NNVs, the phosphoprotein plays multiple roles during the replication process, some of which seem to be common, while others are specific to a virus or a viral family. In all these viruses, P forms a two-subunit RNA-dependent RNA polymerase with the L protein, in which P acts as a noncatalytic cofactor, positioning the polymerase on its template (14). The genomic RNA of the rhabdoviruses, like that of all NNVs, is entirely coated by the viral N, and it is the N-RNA complex rather than the naked RNA that serves as a template for the polymerase (15). The replication of the viral genome thus requires a continuous supply of N, and a second common function of P is to act as a chaperone of the nascent nucleoprotein, preventing the automatic assembly of N in the absence of viral RNA and the encapsidation of cellular RNA (16, 17).

VSV P and RAV P form nonglobular dimers (18). Limited proteolysis, disorder predictions, and biochemical studies of protein fragments revealed that these proteins are multidomain proteins that consist of two structured domains and two disordered regions (19). A central domain is responsible for the dimerization of the protein, whereas a C-terminal domain binds to the N-RNA template. The structures of both domains of VSV and RAV P, as well as the C-terminal domain of Mokola virus (MOKV), were solved recently (20–24). Despite a lack of sequence conservation, the structures of the three C-terminal domains are similar (24, 25), suggesting that they are homologous. Conversely, the structures of the central domains of VSV and RAV are so different that

Received 11 April 2013 **Accepted** 14 June 2013
Published ahead of print 19 June 2013
Address correspondence to Marc Jamin, jamin@embl.fr.

doi:10.1128/JVI.00999-13

it is impossible to conclude that they have a common ancestry (22). The N-terminal region of the protein is globally disordered, while a flexible linker connects the central and C-terminal domains, making the protein highly flexible and structurally heterogeneous. A recent integrative structural characterization of the phosphoprotein of VSV combining the atomic structures of the domains with nuclear magnetic resonance spectroscopy and small-angle X-ray scattering data led to a representation of the protein in the form of an equilibrium ensemble of interconverting conformers rather than a unique structure (26). Disorder predictions and structural studies showed that a similar modular organization is conserved among P proteins from the *Paramyxoviridae* (27, 28), *Bornaviridae* (29), and *Filoviridae* (30).

Since no structural information was available for any *Cytorhabdovirus*, we used a meta-prediction of disorder to localize the disordered and structured regions in the sequence of LNYV P. The disorder score (*D* score) provides a consensus from 16 different web servers that predict the localization of disordered regions from the amino acid sequence of a protein (19, 24). Consequently, for a multidomain protein, this meta-prediction of disorder allows localization of the boundaries between ordered and disordered regions. On the basis of this analysis, we cloned, expressed, and purified the C-terminal domain of the LNYV P protein. Its crystal structure revealed a fold that is different from the folds of VSV and RAV and from the folds of paramyxo- and filoviruses but that suggests common ancestry.

MATERIALS AND METHODS

Amino acid sequence analysis, cloning, production, and purification of the C-terminal domain of LNYV P. The location of disordered regions within LNYV P was predicted by submitting its amino acid sequence to 16 different algorithms accessible through web servers and by calculating a consensus prediction, as described previously (19, 24). On the basis of this prediction, a synthetic gene (GeneArt) optimized for expression in *Escherichia coli* and encoding residues 230 to 300 of LNYV P (71 residues) was cloned between the NcoI and BamHI restriction sites of a pET28 vector, such that the construct includes two additional N-terminal residues (Met-Ala) and a C-terminal His_6 tag with a Leu-Glu linker. The construction was verified by DNA sequencing. The protein was expressed in *E. coli* BL21(DE3-RIL). Cells were grown in LB medium at 37°C until the optical density at 600 nm reached a value of 0.6. Then, protein expression was induced by adding 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside), and cells were grown for 3 h. Protein with a selenomethionine (SeMet) substitution was produced by growing the bacterial cells in minimal medium and by adding SeMet before induction of the heterologous protein expression. The cells were harvested and suspended in 20 mM Tris-HCl buffer at pH 7.5 containing 150 mM NaCl and EDTA-free protease inhibitor cocktail (Roche) (buffer A). The cells were broken by sonication, and the soluble fraction was loaded onto a Ni^{2+} column equilibrated in buffer A. The column was washed with buffer A containing 20 mM imidazole, and the protein was then eluted with buffer A containing 300 mM imidazole. The protein was further purified by gel filtration with a Superdex S75 column (GE Healthcare) equilibrated in buffer A. The protein preparations were checked by SDS-PAGE.

SEC combined with detection by MALLS and refractometry. Size-exclusion chromatography (SEC) was performed with a Superdex S75 column (GE Healthcare) equilibrated in 20 mM Tris-HCl, pH 7.5, 150 mM NaCl. Separations were performed at 20°C with a flow rate of 0.5 ml · min^{-1}. Fifty microliters of a protein solution at a concentration of 7 mg · ml^{-1} was injected. Online multiangle laser light scattering (MALLS) detection was performed with a DAWN-EOS detector (Wyatt Technology Corp., Santa Barbara, CA) using a laser emitting at 690 nm. The protein concentration was measured online by the use of refractive index measurements, an RI2000 detector (Schambeck SFD), and a refractive index increment *dn*/*dc*, the differential increment of refractive index per differential increment of concentration, of 0.185 ml · g^{-1}. Data were analyzed, and weight-average molecular weights (M_ws) were calculated using the software ASTRA V (Wyatt Technology Corp., Santa Barbara, CA) as described previously (18). For size determination, the column was calibrated with proteins of known Stokes radius (R_S) (31).

CD spectroscopy. Far-UV circular dichroism (CD) spectra were recorded at 20°C on a Jasco model J-810 CD spectropolarimeter equipped with a Peltier temperature controller. The LNYV P C-terminal domain (P_{CTD}) was in 20 mM Tris-HCl, pH 7.5, containing 150 mM NaCl. Spectra were measured in a cuvette with a path length of 1 mm. The temperature-induced unfolding and refolding were recorded at 222 nm in a cuvette with a path length of 10 mm by varying the temperature between 25 and 60°C using a scan rate of 60°C/h.

Crystallization, data collection, and structure determination and refinement. Crystallization conditions were screened by the hanging-drop vapor diffusion method using a PixSys4200 Cartesian robot (high-throughput crystallization laboratory at EMBL, Grenoble, France). The screen was performed by combining 0.1 μl of protein solution at 8 mg · ml^{-1} in buffer A with 0.1 μl of Hampton crystal screen solutions.

The selenomethionine derivative of LNYV P_{CTD} was crystallized in a solution of 20 mM sodium cacodylate buffer at pH 6.5 containing 200 mM magnesium acetate and 15% 2-methyl-2,4-pentanediol (MPD). Crystals were frozen in the same solution at a final concentration of 30% MPD as the cryoprotectant.

Diffraction data were collected at Se-K edge on the ID14-4 beam line (European Synchrotron Radiation Facility [ESRF], Grenoble, France). The crystals diffracted to 2.0-Å resolution and belonged to space group $P4_122$. Data were integrated and scaled using the XDS program (82). Two selenium sites and initial phases were found by the single-wavelength anomalous diffraction (SAD) method using the ShelX-C-D-E pipeline (83), and the software Resolve (84) was used to automatically trace a model in the resulting electron density. The model was improved by hand and finally refined with the REFMAC5 program (85). The quality of the model was checked with the PROCHECK program (32). Data collection and refinement statistics are summarized in Table 1.

Structure analysis and comparisons. The CLUSTALW2 server was used for sequence alignments (33). The structure homology program (SHP) was used for structural alignment and for constructing the structure-based phylogenetic tree (34, 35). The electrostatic maps were calculated using the Delphi program (36). The conservation score was calculated with the AL2CO program (37). Figures of the structure were prepared with the PyMOL program (38).

Protein structure accession number. Coordinates of the protein have been deposited in the Research Collaboratory for Structural Bioinformatics Protein Data Bank (PDB) with the accession number 3T4R.

RESULTS

Meta-prediction of disorder. The structural characterization of the phosphoprotein from different NNV families revealed a modular architecture made of folded domains concatenated with disordered regions (27). Multiple-sequence alignments, however, showed no significant sequence conservation between the amino acid sequence of LNYV P and the sequences of the P proteins from other *Rhabdoviridae* or NNVs (39). A consensus from multiple disorder predictions was calculated in the form of a *D* score, using a method previously described (19, 24). In these studies, plotting the *D* score as a function of residue number allowed localization of the boundaries of the oligomerization domain (P central domain, P_{CED}) and N-RNA binding C-terminal domain (P_{CTD}) in the P proteins from vesiculoviruses (VSV) and lyssaviruses (RAV) (19, 24). The *D* score calculated for the amino acid sequence of LNVY P predicted the presence of two intrinsically disordered regions

TABLE 1 Data collection and refinement statistics

Parameter[a]	Value for crystal 1[b]
Data collection statistics	
Space group	$P4_122$
Unit cell dimensions	
a, *b*, *c* (Å)	43.45, 43.45, 89.71
α, β, γ (°)	90.0, 90.0, 90.0
Resolution (Å)	43.28–2.00 (2.16–2.00)
R_{sym} (%)	4.3 (26.6)
Avg $I/\sigma<I>$	16.05(4.03)
Completeness (%)	98.9 (100)
Redundancy	3.76 (3.80)
Refinement statistics	
Resolution (Å)	43.28–2.00 (2.16–2.00)
No. of reflections	47,380 (7,366)
R_{work}/R_{free} (%)	22.6/29.0
No. of atoms	
Protein	610
Water	22
RMSD	
Bond lengths (Å)	0.031
Bond angles (°)	2.095
Avg B factor (Å^2)	
Protein	54.8
Water	58.0
Ramachandran plot (%)	
Favored regions	90.8
Allowed regions	98.2

[a] R_{sym}, linear R factor used to estimate the precision with which diffraction data are measured; *I*, intensity of a reflection.
[b] Values in parentheses are for the highest-resolution shell.

(IDRs) with contiguous *D*-score values lower than 0.5 (N-terminal IDR, amino acids [aa] 1 to 63; C-terminal IDR, aa 184 to 229) and two folded regions with contiguous *D*-score values higher than 0.5, a central bipartite domain (P protein central bipartite domain, aa 64 to 183) and a C-terminal domain (P_{CTD}, aa 230 to 300) (Fig. 1). This modular organization, reminiscent of that previously observed for other NNV viruses, suggested that the predicted C-terminal domain of LNYV P corresponded to the N-RNA binding domain of rhabdovirus P proteins (19, 23, 24).

Expression and purification of LNYV P_{CTD}. On the basis of the boundaries defined by the *D*-score analysis, a synthetic cDNA corresponding to residues 230 to 300 of LNYV P was cloned and expressed in the *Escherichia coli* BL21(DE3-RIL) strain as a His tag fusion protein. The protein was purified by nickel-affinity chromatography and size-exclusion chromatography. In solution, P_{CTD} behaves as a monomer with a molecular mass of 9.2 ± 0.5 kDa, determined by SEC-MALLS, which is in good agreement with the mass of 9,198 Da calculated from the amino acid sequence (including nonviral amino acids from the N- and C-terminal extremities) (Fig. 2A). The Stokes radius of 1.7 ± 0.1 nm is

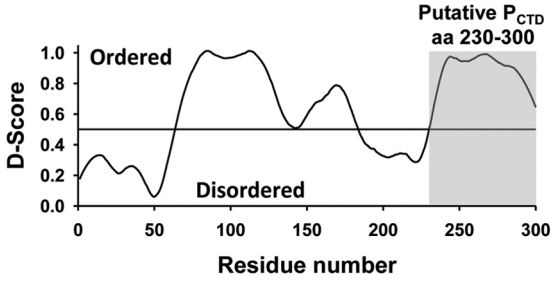

FIG 1 Predicting boundaries between structured and disordered regions for LNYV P. The *D* score as a function of residue number is calculated as a consensus of different predictions, as described in Materials and Methods. Regions of the protein with contiguous *D*-score values above 0.5 are predicted to be structured, whereas regions with *D*-score values below 0.5 are predicted to be disordered (19). The gray shaded area shows the region of the protein that has been selected, cloned, and characterized in this work.

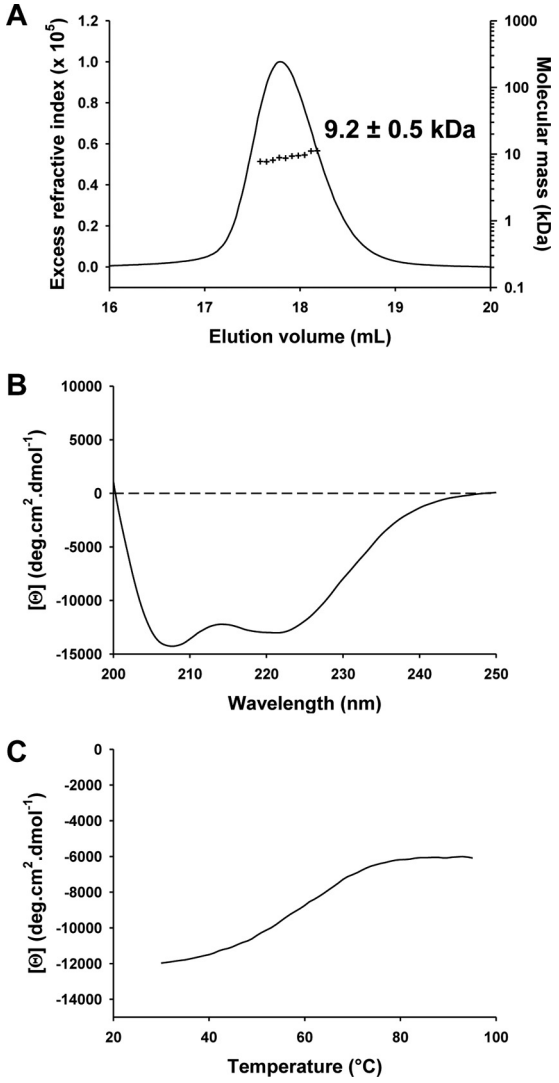

FIG 2 Solution properties of LNYV C-terminal domain. The experiments were performed in a 20 mM Tris-HCl buffer at pH 7.5 containing 150 mM NaCl at 20°C. (A) Molecular mass determined by size-exclusion chromatography combined with detection by multiangle laser light scattering and refractometry (SEC-MALLS-RI). Fifty microliters of a protein sample at 7 mg · ml^{-1} was injected on a Superdex S75 column. The line shows the SEC elution profile as monitored by refractometry. The crosses show the molecular mass calculated from light scattering and refractometry data. (B) Secondary structure by far-UV circular dichroism. The spectrum was measured with a protein sample at a concentration of 2 mg · ml^{-1} in a 1-mm cuvette. (C) Thermal unfolding monitored by far-UV circular dichroism at 222 nm. The experiment was performed with a protein sample at a concentration of 0.2 mg · ml^{-1} in a 10-mm cuvette under continuous stirring. [Θ], molar ellipticity per residue. The temperature was raised at a speed of 1°C · min^{-1}.

close to that expected for a globular protein of this molecular mass (1.6 nm) (31). The far-UV circular dichroism spectrum with two minima at 208 and 222 nm indicated an α-helix-rich protein (Fig. 2B). The reversible thermal unfolding exhibited a sigmoid transition typical of a well-folded protein with a hydrophobic core (Fig. 2C). These results clearly demonstrated that LNYV P_{CTD} consists of an autonomous structural unit that adopts a globular fold with a large content of α-helical structure even in the absence of the remaining part of the protein.

Crystal structure of LNYV P_{CTD}. A selenomethionine derivative of the LNYV P_{CTD} was crystallized at 20°C in 20 mM sodium cacodylate buffer at pH 6.5 containing 200 mM magnesium acetate and 15% MPD. Diffraction data on frozen crystals were collected at the Se-K edge on beam line ID14-4 (ESRF, Grenoble, France). The crystal diffracted to a 2.0-Å resolution and belonged to space group $P4_122$ with one monomer in the asymmetric unit. The phases were solved by the SAD method using the ShelX-C-D-E pipeline (40). A preliminary model including residues 230 to 300 of LNYV and the first amino acid (Leu) of the C-terminal His tag was automatically generated in the electron density with the software Resolve (41) and was refined with the REFMAC5 program (42). The quality of the model was evaluated with the PROCHEK program (32) (Table 1). The high B-factor values associated with numerous residues in the structure and the absence of visible density for eight of the nonviral residues that are part of our construct likely contribute to the discrepancy found between the quality of the data (resolution of 2 Å) and the quality of the derived model (R_{work} = 22.6%; R_{free} = 29.0%) (Table 1).

The C-terminal domain of LNYV P comprises five α helices, named α_1 to α_5 (Fig. 3A). The overall structure of P_{CTD} is made of a globular region formed by helices α_1 to α_4 and by the N-terminal part of helix α_5 and of a stalk formed by the C-terminal part of helix α_5. The globular region has one flat face formed by helices α_4 and α_5 and one rounded face formed by the three remaining helices. Side chains from residues in helices α_1 to α_4 and residues in the N-terminal moiety of helix α_5 form a hydrophobic core. A hydrophobic groove including a hole is localized on the protein surface between helices α_1, α_2, and α_5 (Fig. 3B, hydrophobic groove 1). It includes Trp244, Tyr252, Ile256, Phe259, Leu260, and Leu292. A second deep groove at the top of the domain is lined with Leu238, Leu242, Leu266, and Leu273 (Fig. 3B, hydrophobic groove 2). Analysis of the distribution of electrostatic potentials (Fig. 3C) showed the presence of two patches of negative charges on opposite faces of the molecule. Asp263, Asp264, Glu284, and Glu287 form an elongated negative patch on the flat face (negative patch 1), whereas Asp243, Glu245, Glu248, Glu276, and Glu279 (negative patch 2) line along a ridge on the round face of P_{CTD}.

Structural comparison with other P_{CTD}s. The three-dimensional structure of the corresponding C-terminal domain of P has been solved previously for three rhabdoviruses (20, 23, 24), three paramyxoviruses (43–45), and the VP35 protein of a filovirus (46, 47). In the absence of sequence conservation between LNYV P and these other P proteins, we performed pairwise alignment of the LNYV P_{CTD} structure with the P_{CTD} structures from the other viruses using the PDBeFold server (48) and the TM-align server (49). These programs use the optimization of empirical scores for aligning structures. These scores, the quality function score (*Q* score) associated with the probability *P* score used by PDBeFold and the template-modeling score (TM score) used by TM-align, account for both alignment length and the distance between aligned residues and, therefore, are less sensitive to local structural variations than the root mean square deviation (RMSD) alone (Table 2). The value of these scores increases from 0 to 1 with increasing structural similarity. Typically, a high *Q*-score value associated with a *P* value larger than 3 or a TM value larger than 0.5 indicates that there is a similarity in the fold of the aligned molecules (48, 49). Both procedures yielded values below the limits considered to indicate structural homology (Table 2). The *Q*-score values were between 0.10 and 0.25, but the associated *P* values were low, indicating a low significance. The TM-score values approached the 0.5 limit but remained slightly lower. In conclusion, these alignment procedures suggest that the structure of the LNYV P_{CTD} domain is different from that of the P_{CTD} domains of other NNVs.

Search of PDB. Searches of PDB using the DALI server (50) or PDBeFold server (48) yielded only hits that did not meet the significance criteria set by default in these processes and failed to recognize the C-terminal domains of other NNV phosphoproteins as relatives of the LNYV P_{CTD}.

DISCUSSION

A common modular architecture for NNV phosphoproteins. All NNVs share a similar organization of their genome, similar mechanisms for transcription and replication in the host cell, and similar functions for the different viral proteins in the replication process, arguing for a descent from a common ancestor. However, the high rate of mutations due to the absence of proofreading in RNA replication produces large divergences in protein lengths and sequences. For some proteins, homology between the different NNV families remains detectable at the level of the amino acid sequence, allowing the construction of phylogenetic trees from multiple-sequence alignments (3, 51). Structural comparison of the nucleoproteins and matrix proteins clearly reveals conservation of protein folds (52, 53). For the phosphoprotein, however, the length of the protein varies greatly between viruses, and sequence similarity is generally undetectable beyond the family level and even sometimes beyond the genus level (3, 39, 54).

Predictions and experimental studies revealed that the phosphoproteins from different NNV families share a modular architecture with disordered regions alternating with folded domains (19, 27, 55, 56). Disordered regions and folded domains playing similar roles in the replication and transcription processes are positioned in the same order along the polypeptide chain. The protein can be grossly divided into two parts. The N-terminal part is globally disordered and in some cases contains transiently populated helices in regions involved in binding viral or cellular partners. In particular, this N-terminal region contains a binding site for the RNA-free nucleoprotein (N^0) conserved in the amino acid sequence (39). For VSV, the *D* score predicted this region to be structured (19), but it is globally disordered in isolation and forms a stable helix only upon binding to N^0 (57, 58). The C-terminal part contains a homo-oligomerization domain and the C-terminal N-RNA binding domain (P_{CTD}) connected by a flexible linker of variable length. The P protein of the *Paramyxoviridae* forms tetramers (28, 59), whereas that of the *Rhabdoviridae* forms dimers (18, 21, 22). The *D* score calculated from the amino acid sequence of LNYV (Fig. 1) supports a modular organization similar to that of the other NNVs and predicts the presence of a long N-terminal disordered region (aa 1 to 63) and a central folded region (aa 64 to 184) separated from P_{CTD} (aa 230 to 300) by a

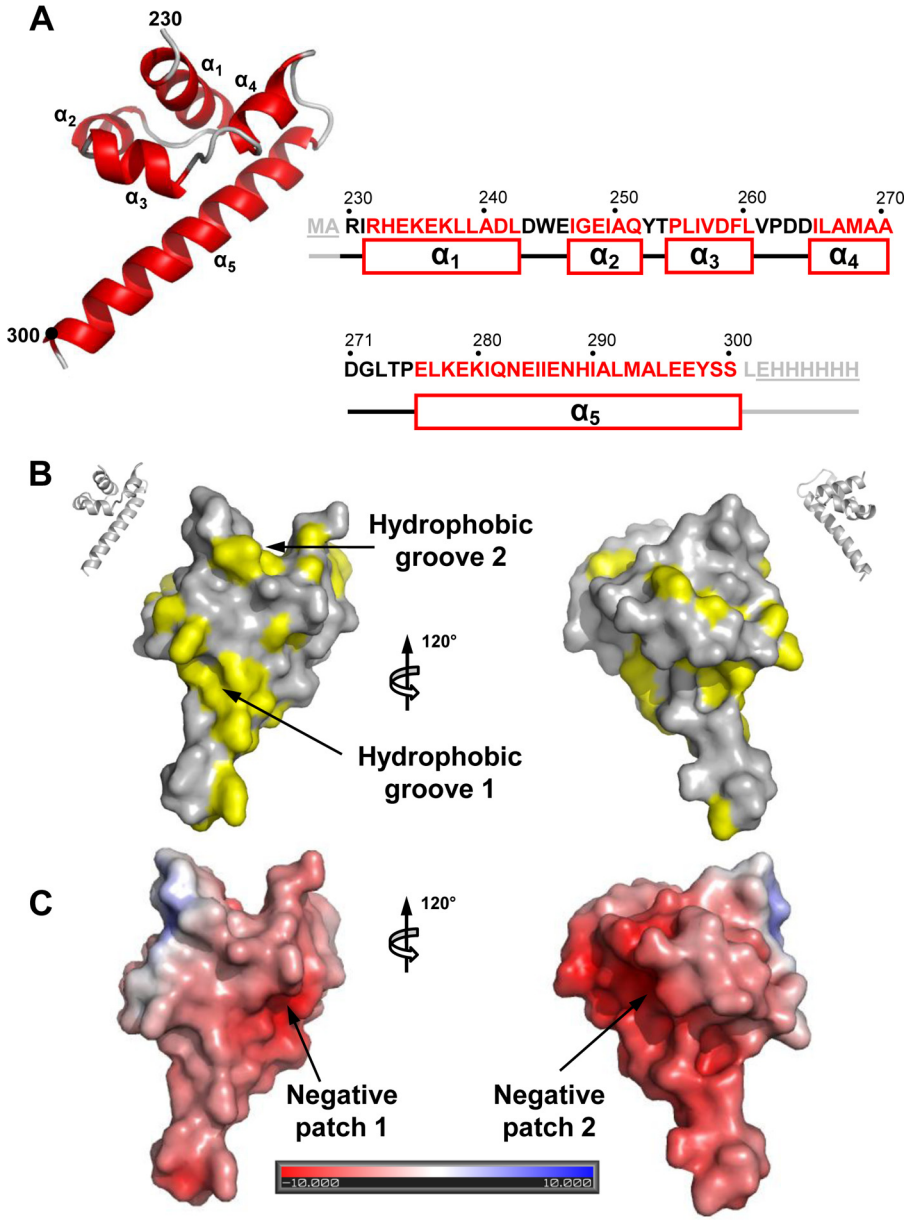

FIG 3 Crystal structure of LNYV P_{CTD}. (A) Schematic representation of LNYV P_{CTD}. Helices are numbered from α_1 to α_5. The numbers indicate the positions of the N- and C-terminal residues. The figure was drawn with the PyMOL program. The positions of the α helices along the amino acid sequence are shown. The underlined residues are not visible in the crystal. The residues of LNYV are shown in black or red, and the nonviral residues added in the construct are colored in gray. (B) Hydrophobic surface map. The distribution of hydrophobic residues on the surface is shown; yellow indicates hydrophobic sidechains. (C) Electrostatic surface potential. The surface potential calculated with the Delphi program is plotted on the solvent-accessible surface of the protein from red (negatively charged regions in electrostatic potential units, where k_B is Boltzmann's constant, T is the temperature, and e is the electron charge; $-10\ k_BTe^{-1}$) to blue (positively charged regions; $+10\ k_BTe^{-1}$).

disordered linker (aa 185 to 229). The predicted central domain of LNYV may correspond to the homo-oligomerization domain.

Evolutionary relationships within NNV families. The scores obtained for pairwise structural alignments of the P_{CTD}s from the *Rhabdoviridae*, *Filoviridae*, and *Paramyxovirinae* are below the threshold commonly accepted for validating homology. Even the structural comparisons between LNYV P_{CTD} and the corresponding domain from vesiculoviruses or lyssaviruses show no evidence for structural conservation. Nevertheless, because different evidence supports a divergent evolution of these viral families (54, 60) and because in all NNV P proteins this domain is located at the C-terminal end and fulfills the same function of binding to the nucleocapsid, it is tempting to speculate that they diverged from a common ancestor. As such, structural comparisons can provide clues about the mechanisms by which this domain has evolved. Indeed, the pairwise alignments of the LNYV P_{CTD} structure with the P_{CTD} structures from the other viruses showed a remarkable conservation of the topology of the backbone and of secondary structure elements (Fig. 4). Helices α_1, α_2, α_3, α_4, and α_5 of LNYV P_{CTD} are aligned with helices α_1, η_1 (3^{10} helix), α_2, α_3, and α_4 of

TABLE 2 Structural alignments[a]

Virus	PDB accession no.	N_{target}	PDBeFold				TM-align		
			Q	P	RMSD (Å)	N_{align}	TM	RMSD (Å)	N_{align}
RAV	1VYI	111	0.18	0.00	3.24	55	0.43	3.31	50
MOKV	2WZL	107	0.15	0.06	3.15	50	0.47	3.32	58
VSV	2K47	73	0.13	0.03	3.37	39	0.36	3.26	72
MEV	10KS	53	0.25	0.00	3.41	47	0.41	3.53	47
SEV	1R4G	53	0.20	0.02	3.42	42	0.40	3.01	45
MUV	3BBZ	48	0.19	0.00	3.94	42	0.40	2.98	48
EBOV	3FKE	123	0.10	0.00	3.26	44	0.34	3.84	43
WASP	1EJ5	107	0.18	0.14	3.73	59	0.47	3.24	61

[a] LNYV P_{CTD}, used as the query, contains 71 residues. N_{target}, number of residues in the target structure; N_{align}, number of residues aligned in the best three-dimensional superposition. RMSD is calculated between C_α atoms of matched residues in the best three-dimensional superposition according to the following equation:

$$\mathrm{RMSD} = \sqrt{\frac{1}{N_{res}}\sum_{i=1}^{N_{res}} d_i}$$

where d_i is the distance between N pairs of equivalent C_α atoms, i is the residue number, and N_{res} is the number of residues. The Q score evaluates the quality of the C_α atom alignment and takes both RMSD and alignment length into account according to the following equation:

$$\frac{N_{align1}\,N_{align2}}{\left[1 + (\mathrm{RMSD}/R_0)^2\right]N_{res1}\,N_{res2}}$$

where N_{align1} and N_{align2} are the number of aligned residues in the query and target structures, respectively; N_{res1} and N_{res2} are the number of residues in each protein; and R_0 is an empirical parameter set equal to 3 Å. Q takes values between 0 and 1, and identical structures have a score of 1. The P-score parameter represents the minus logarithm of the P value, where the P value measures the probability of achieving a match of the same or better quality as that obtained by chance. This P score was calibrated using a set of 700 structures, and P scores less than 3 indicate statistically insignificant matches. The TM score evaluates the quality of the C_α alignment, which takes both RMSD and alignment length into account according to the following equation:

$$\mathrm{TM} = \max\left\{\frac{1}{N_{res}}\sum_{i}^{N} 1/[1 + (d_i/d_0)^2]\right\}$$

where max is maximum and d_0 is an empirical parameter depending on the length of the protein:

$$d_0 = 1.24\sqrt[3]{N_{res} - 15} - 18.$$

TM takes values between 0 and 1. TM values of <0.17 correspond to random similarities, whereas TM values of >0.5 indicate an equivalent fold. MEV, measles virus; SEV, Sendai virus; MUV, mumps virus; WASP, Wiskott–Aldrich syndrome protein.

RAV and MOKV P_{CTD} (Fig. 4B), respectively, suggesting that the overall topological organization of these secondary structure elements is conserved among rhabdoviruses. Previously, we had shown that the topology of P_{CTD} is also conserved between RAV and VSV but that the latter lacks helices α_4 and α_6 of the RAV protein (24). In accordance, helices α_1, α_2, α_3, and α_5 of LNYV P_{CTD} are matched with α_1, η_1, α_2, and α_4 of VSV, respectively (Fig. 4C). These alignments suggest that the P_{CTD} domain of rhabdoviruses may have conserved the backbone topology of an ancient fold and may have evolved by acquiring or losing secondary structure elements.

The C-terminal domain of paramyxovirus P proteins consists of an antiparallel triple-helix bundle that seems to be well conserved within this family and is different from the structure of rhabdovirus P_{CTD}. Although it may be ambiguous to align a domain comprising only three consecutive helices, the four helices of LNYV P_{CTD} can be aligned with the three helices of measles virus, mumps virus, and Sendai virus P_{CTD} (Fig. 4D). Helices α_2 and α_5 of LNYV P_{CTD} match helices α_1 and α_3 of the paramyxoviruses, respectively. The pair of short helices of LNYV, α_3 and α_4, then corresponds to helix α_2 of the paramyxoviruses, although only one of these helices may be conserved and the other may correspond to an additional structural element. These structural alignments suggest some similarities in the general topology of the polypeptide backbone and in the arrangement of secondary structure elements, as already suggested (24, 25, 39), even if our structural alignment of P_{CTD} from the *Rhabdoviridae* and *Paramyxoviridae* is different from that previously reported (25). Similarly, the LNYV P_{CTD} structure can also be aligned with the C-terminal domain of the Ebola virus (EBOV) VP35 protein. This domain is composed of an α-helical subdomain and a β-sheet subdomain, and the five helices of LNYV can be aligned with the four helices of the helical subdomain of EBOV P_{CTD}, again, with the short helices α_3 and α_4 of LNYV matching the longer α_3 helix of EBOV (Fig. 4E).

To translate these structural similarities into evolutionary distances and to build a phylogenetic tree based on the P_{CTD} structures, we used the structure homology program (34, 35). According to this procedure, pairwise alignments were performed, and probabilities of equivalence were calculated for every residue of both structures by comparing the positions of their C_α atoms and the local shape of the polypeptide chains around each residue. A global parameter characteristic of the similarity between the two structures was then computed as the sum of the probabilities for all residues that have a structural equivalent in both structures. This parameter was further converted into an evolutionary distance using an empirical logarithmic function established with metrics for sequence alignments, and these distances were then used to build a phylogenetic tree. The structure-based phylogenetic tree for the NNVs obtained from the comparison of their P_{CTD} domains (Fig. 5) is similar to the trees obtained on the basis of the amino acid sequence of the N or L proteins (3, 52). The tree suggests a monophyletic origin for the P_{CTD} domain. It shows that cytorhabdoviruses are as evolutionarily distant from other *Rhabdoviridae* of the genera *Vesiculovirus* and *Lyssavirus* as from the *Filoviridae* and *Paramyxovirinae* and suggests that the C-terminal

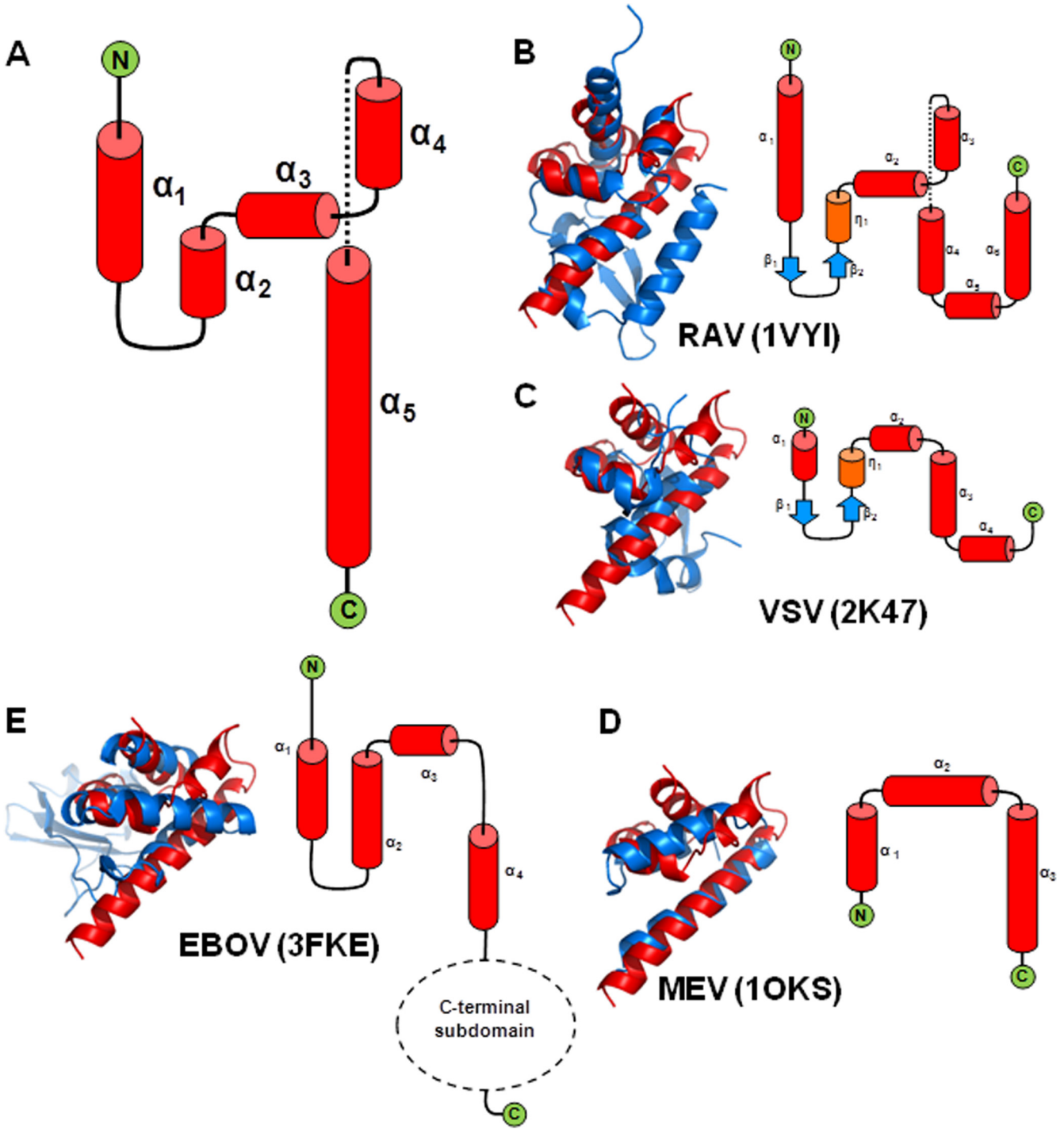

FIG 4 Comparison of the LNYV P_{CTD} structure with the P_{CTD} structures of other NNVs. (A) Topology of LNYV P_{CTD}. (B to E) Pairwise structural alignments. The structure of LNYV was aligned pairwise with the structure of the C-terminal domains of different rhabdoviruses (RAV, VSV, MOKV), paramyxoviruses (measles virus, mumps virus, Sendai virus), and filoviruses (EBOV) (Table 2). Representative superpositions for RAV (B), VSV (C), measles virus (MEV) (D), and EBOV (E) are shown. In each panel, the structure of LNYV P_{CTD} is shown in red and that of the aligned domain is shown in blue. The topology of the compared protein domain is shown by the side.

domain of P has evolved after the NNVs diverged into the different families. Most plant rhabdoviruses are dependent for their transmission on insects, in which they also multiply, and are thereby influenced in their evolution by replication in their vectors (61). In this context, it is notable that, within the *Rhabdoviridae*, LNYV is slightly more closely related to VSV, which is also an arbovirus, than to the lyssaviruses (RAV and MOKV).

In the hypothesis that P from the different NNV families

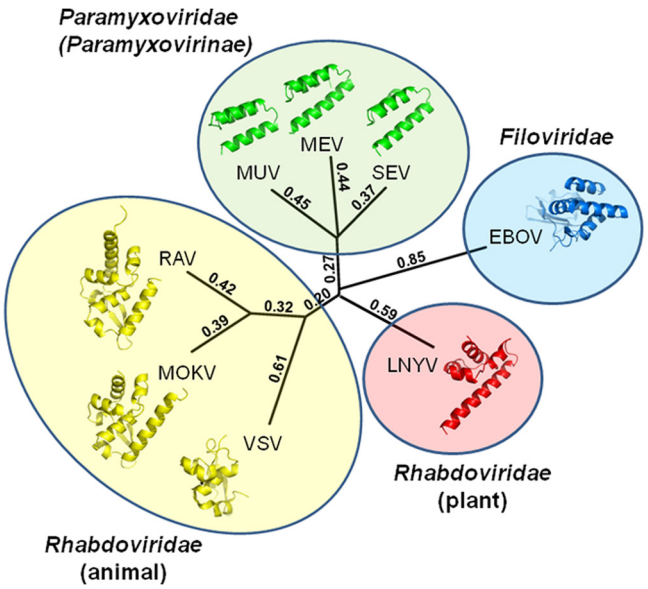

FIG 5 Structure-based phylogenetic tree showing the three NNV families. Using the structure homology program for superposition, the tree shows the P_{CTD} structure of the following viruses: vesicular stomatitis virus (VSV; PDB accession no. 2K47), rabies virus (RAV; PDB accession no. 1VYI), Mokola virus (MOKV; PDB accession no. 2WZL), lettuce necrotic yellows virus (LNYV; PDB accession no. 3T4R), Ebola virus (EBOV; PDB accession no. 3FKE), measles virus (MEV; PDB accession no. 1OKS), mumps virus (MUV; PDB accession no. 3BBZ), and Sendai virus (SEV; PDB accession no. 1R4G). The length of the branches is scaled to structural differences, such that the distance between two species is the sum of the length of all branches connecting them. The numbers relative to the distance derived from this analysis are indicated above each branch.

evolved from a common ancestor, we can ask the question of why its C-terminal domain varies so much in sequence and length. The C-terminal domain plays multiple roles in the different viruses, but its common function is to mediate the binding of P to the nucleocapsid (62, 63), an interaction that seems to be essential for the processivity of the viral polymerase (64, 65). The large variation in P_{CTD} sequences can be explained by weak selective pressure acting on the sequence, as the only functional requirement is the conservation of an interaction with the nucleocapsid that is sufficiently strong for attaching and positioning the polymerase onto its template. Such selective pressure is much less stringent than it is, for example, in the case of an enzyme active site, and therefore, the P_{CTD} sequence is likely to tolerate extensive sequence variations. In addition, the topological complexity of the polypeptide chain is low, as judged by values of the contact order parameter of less than 0.2 (66), and therefore also suggests low selective pressure on the amino acid sequence from folding. It is worth pointing out that this variability extends even further to the mechanism of complex formation between P and the N-RNA template. P_{CTD} of the rhabdoviruses is a well-folded domain that binds to the top of the C-terminal domain of one N protomer, with the flexible N C-terminal loop of the same protomer and of the adjacent one pinching P_{CTD} on two sides (62, 63). In the *Filoviridae*, the subdomain of P_{CTD} that is directly involved in the interaction with the nucleocapsid has been localized, but no structural information about the formation of the complex is available. In the subfamily *Paramyxovirinae*, P_{CTD} binds to the C-terminal disordered tail of N that protrudes from the nucleocapsid (67), inducing the formation of a helix in this disordered region of N upon formation of the complex (68–70). In the subfamily *Pneumovirinae*, the region corresponding to P_{CTD} is a short, disordered segment that binds to the surface of the N-terminal domain of N (71, 72). In the latter case, it is even more difficult to detect structural homology. One may speculate that, for the *Pneumovirinae*, the C-terminal region of P has been extensively reduced to a minimal segment which remains capable of binding to N, though at a different site on the surface of N.

In addition, the structural comparison of P_{CTD} from the different NNV families suggests that the variation in the length of this domain results not only from mutations but also from the extension or reduction of secondary structure elements and potentially also from the acquisition or loss of some secondary structure elements. In the *Paramyxoviridae*, the only known function of P_{CTD} is to attach P to the nucleocapsid, whereas in the *Filoviridae* and *Rhabdoviridae*, P_{CTD} is involved in other functions. We can therefore speculate that the larger size of the C-terminal domain is correlated with its implication in other functions. In EBOV P_{CTD} (VP35), two entities performing different functions are clearly distinguishable in the structure (46): a four-helix subdomain that contains the structurally homologous helices of the *Paramyxovirinae* and interacts with the nucleocapsid (73) and a C-terminal subdomain made of a β sheet, an α helix, and a type II polyproline helix involved in double-stranded RNA binding and in evading the host immune system (46, 47). In animal rhabdoviruses (VSV and RAV), the functional regions are not as clearly separated from each other (23, 24). VSV P_{CTD} interacts with L (74), is phosphorylated by a cellular kinase (75), and regulates viral transcription and replication (76). RAV P_{CTD} is larger than VSV P_{CTD} and is involved in supplementary activities. In addition to its role in the fixation of P to the nucleocapsid and in the regulation of viral transcription and replication through phosphorylation by cellular kinases (77, 78), RAV P_{CTD} is also involved in the evasion of the host immune system through interactions with different cellular factors (79–81) and in the regulation of the nucleocytoplasmic transport of P (77). In conclusion, the P protein may be less conserved than other NNV proteins because it has served as a platform for acquiring new functions through the recruitment of additional structural modules and thereby for allowing viruses to adapt to their different host cell environments.

ACKNOWLEDGMENTS

This work was supported by grants from the French ANR (ANR-07-001-01 [ANRAGE], ANR-12-BSV8-0025-01). E.A.R. was supported by a postdoctoral fellowship from the ANR. N.M. was supported by a fellowship from the Région Rhone-Alpes.

We thank the Partnership for Structural Biology for the excellent structural biology environment.

REFERENCES

1. **Jackson AO, Dietzgen RG, Goodin MM, Bragg JN, Deng M.** 2005. Biology of plant rhabdoviruses. Annu. Rev. Phytopathol. **43:**623–660.
2. **Kormelink R, Garcia ML, Goodin M, Sasaya T, Haenni AL.** 2011. Negative-strand RNA viruses: the plant-infecting counterparts. Virus Res. **162:**184–202.
3. **Dietzgen RG, Callisher CH, Kurath G, Kuzmin IV, Rodriguez LL, Stone DM, Tesh RB, Tordo N, Walker PJ.** 2011. Family *Rhabdoviridae*. *In* King AM, Adams MJ, Carstens EB, Lefkowitz EJ (ed), Virus taxonomy, VIIIth Report of the ICTV. Elsevier Academic Press, Oxford, United Kingdom.
4. **Stubbs LL, Grogan RG.** 1963. Lettuce necrotic yellows virus. Nature **197:**1229.

5. **Heim F, Lot H, Delecolle B, Bassler A, Krczal G, Wetzel T.** 2008. Complete nucleotide sequence of a putative new cytorhabdovirus infecting lettuce. Arch. Virol. **153:**81–92.
6. **Martin KM, Dietzgen RG, Wang R, Goodin MM.** 2012. Lettuce necrotic yellows cytorhabdovirus protein localization and interaction map, and comparison with nucleorhabdoviruses. J. Gen. Virol. **93:**906–914.
7. **Wolanski BS, Chambers TC.** 1971. The multiplication of lettuce necrotic yellows virus. Virology **44:**582–591.
8. **Dietzgen RG, Callaghan B, Wetzel T, Dale JL.** 2006. Completion of the genome sequence of lettuce necrotic yellows virus, type species of the genus Cytorhabdovirus. Virus Res. **118:**16–22.
9. **Callaghan B, Dietzgen RG.** 2005. Nucleocapsid gene variability reveals two subgroups of lettuce necrotic yellows virus. Arch. Virol. **150:**1661–1667.
10. **Dietzgen RG, Francki RI.** 1988. Analysis of lettuce necrotic yellows virus structural proteins with monoclonal antibodies and concanavalin A. Virology **166:**486–494.
11. **Wetzel T, Dietzgen RG, Dale JL.** 1994. Genomic organization of lettuce necrotic yellows rhabdovirus. Virology **200:**401–412.
12. **Wetzel T, Dietzgen RG, Geering AD, Dale JL.** 1994. Analysis of the nucleocapsid gene of lettuce necrotic yellows rhabdovirus. Virology **202:**1054–1057.
13. **Ivanov I, Yabukarski F, Ruigrok RW, Jamin M.** 2011. Structural insights into the rhabdovirus transcription/replication complex. Virus Res. **162:**126–137.
14. **Emerson SU, Yu Y.** 1975. Both NS and L proteins are required for in vitro RNA synthesis by vesicular stomatitis virus. J. Virol. **15:**1348–1356.
15. **Arnheiter H, Davis NL, Wertz G, Schubert M, Lazzarini RA.** 1985. Role of the nucleocapsid protein in regulating vesicular stomatitis virus RNA synthesis. Cell **41:**259–267.
16. **Masters PS, Banerjee AK.** 1988. Complex formation with vesicular stomatitis virus phosphoprotein NS prevents binding of nucleocapsid protein N to nonspecific RNA. J. Virol. **62:**2658–2664.
17. **Peluso RW, Moyer SA.** 1988. Viral proteins required for the in vitro replication of vesicular stomatitis virus defective interfering particle genome RNA. Virology **162:**369–376.
18. **Gérard FCA, Ribeiro E, Albertini A, Zaccai G, Ebel C, Ruigrok R, Jamin M.** 2007. Unphosphorylated Rhabdoviridae phosphoproteins form elongated dimers in solution. Biochemistry **46:**10328–10338.
19. **Gérard FCA, Ribeiro EA, Leyrat C, Ivanov I, Blondel D, Longhi S, Ruigrok RWH, Jamin M.** 2009. Modular organization of rabies virus phosphoprotein. J. Mol. Biol. **388:**978–996.
20. **Assenberg R, Delmas O, Ren J, Vidalain PO, Verma A, Larrous F, Graham SC, Tangy F, Grimes JM, Bourhy H.** 2010. The structure of the nucleoprotein binding domain of the Mokola virus phosphoprotein. J. Virol. **84:**1089–1096.
21. **Ding H, Green TJ, Lu S, Luo M.** 2006. Crystal structure of the oligomerization domain of the phosphoprotein of vesicular stomatitis virus. J. Virol. **80:**2808–2814.
22. **Ivanov I, Crepin T, Jamin M, Ruigrok R.** 2010. Structure of the dimerization domain of the rabies virus phosphoprotein. J. Virol. **84:**3707–3710.
23. **Mavrakis M, McCarthy AA, Roche S, Blondel D, Ruigrok RW.** 2004. Structure and function of the C-terminal domain of the polymerase cofactor of rabies virus. J. Mol. Biol. **343:**819–831.
24. **Ribeiro EA, Jr, Favier A, Gerard FC, Leyrat C, Brutscher B, Blondel D, Ruigrok RW, Blackledge M, Jamin M.** 2008. Solution structure of the C-terminal nucleoprotein-RNA binding domain of the vesicular stomatitis virus phosphoprotein. J. Mol. Biol. **382:**525–538.
25. **Delmas O, Assenberg R, Grimes JM, Bourhy H.** 2010. The structure of the nucleoprotein binding domain of lyssavirus phosphoprotein reveals a structural relationship between the N-RNA binding domains of Rhabdoviridae and Paramyxoviridae. RNA Biol. **7:**322–327.
26. **Leyrat C, Schneider R, Ribeiro EA, Jr, Yabukarski F, Yao M, Gerard FC, Jensen MR, Ruigrok RW, Blackledge M, Jamin M.** 2012. Ensemble structure of the modular and flexible full-length vesicular stomatitis virus phosphoprotein. J. Mol. Biol. **423:**182–197.
27. **Karlin D, Ferron F, Canard B, Longhi S.** 2003. Structural disorder and modular organization in Paramyxovirinae N and P. J. Gen. Virol. **84:**3239–3252.
28. **Llorente MT, Garcia-Barreno B, Calero M, Camafeita E, Lopez JA, Longhi S, Ferron F, Varela PF, Melero JA.** 2006. Structural analysis of the human respiratory syncytial virus phosphoprotein: characterization of an alpha-helical domain involved in oligomerization. J. Gen. Virol. **87:**159–169.
29. **Hock M, Kraus I, Schoehn G, Jamin M, Andrei-Selmer C, Garten W, Weissenhorn W.** 2010. RNA induced polymerization of the Borna disease virus nucleoprotein. Virology **397:**64–72.
30. **Leung DW, Prins KC, Basler CF, Amarasinghe GK.** 2010. Ebolavirus VP35 is a multifunctional virulence factor. Virulence **1:**526–531.
31. **Uversky VN.** 1993. Use of fast protein size-exclusion liquid chromatography to study the unfolding of proteins which denature through the molten globule. Biochemistry **32:**13288–13298.
32. **Laskowski RA, Macarthur MW, Moss DS, Thornton JM.** 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26:**283–291.
33. **Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947–2948.
34. **Abrescia NG, Bamford DH, Grimes JM, Stuart DI.** 2012. Structure unifies the viral universe. Annu. Rev. Biochem. **81:**795–822.
35. **Stuart DI, Levine M, Muirhead H, Stammers DK.** 1979. Crystal structure of cat muscle pyruvate kinase at a resolution of 2.6 Å. J. Mol. Biol. **134:**109–142.
36. **Rocchia W, Sridharan S, Nicholls A, Alexov E, Chiabrera A, Honig B.** 2002. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects. J. Comput. Chem. **23:**128–137.
37. **Pei J, Grishin NV.** 2001. AL2CO: calculation of positional conservation in a protein sequence alignment. Bioinformatics **17:**700–712.
38. **DeLano WL.** 2002. The PyMOL molecular graphics system. DeLano Scientific, Palo Alto, CA.
39. **Karlin D, Belshaw R.** 2012. Detecting remote sequence homology in disordered proteins: discovery of conserved motifs in the N-termini of Mononegavirales phosphoproteins. PLoS One **7:**e31719. doi:10.1371/journal.pone.0031719.
40. **Hubschle CB, Sheldrick GM, Dittrich B.** 2011. ShelXle: a Qt graphical user interface for SHELXL. J. Appl. Crystallogr. **44:**1281–1284.
41. **Terwilliger TC.** 2003. Automated main-chain model building by template matching and iterative fragment extension. Acta Crystallogr. D Biol. Crystallogr. **59:**38–44.
42. **Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA.** 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. **67:**355–367.
43. **Blanchard L, Tarbouriech N, Blackledge M, Timmins P, Burmeister WP, Ruigrok RW, Marion D.** 2004. Structure and dynamics of the nucleocapsid-binding domain of the Sendai virus phosphoprotein in solution. Virology **319:**201–211.
44. **Johansson K, Bourhis JM, Campanacci V, Cambillau C, Canard B, Longhi S.** 2003. Crystal structure of the measles virus phosphoprotein domain responsible for the induced folding of the C-terminal domain of the nucleoprotein. J. Biol. Chem. **278:**44567–44573.
45. **Kingston RL, Gay LS, Baase WS, Matthews BW.** 2008. Structure of the nucleocapsid-binding domain from the mumps virus polymerase; an example of protein folding induced by crystallization. J. Mol. Biol. **379:**719–731.
46. **Leung DW, Ginder ND, Fulton DB, Nix J, Basler CF, Honzatko RB, Amarasinghe GK.** 2009. Structure of the Ebola VP35 interferon inhibitory domain. Proc. Natl. Acad. Sci. U. S. A. **106:**411–416.
47. **Leung DW, Prins KC, Borek DM, Farahbakhsh M, Tufariello JM, Ramanan P, Nix JC, Helgeson LA, Otwinowski Z, Honzatko RB, Basler CF, Amarasinghe GK.** 2010. Structural basis for dsRNA recognition and interferon antagonism by Ebola VP35. Nat. Struct. Mol. Biol. **17:**165–172.
48. **Krissinel E, Henrick K.** 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. **60:**2256–2268.
49. **Zhang Y, Skolnick J.** 2005. TM-align: a protein structure alignment algorithm based on the TM-score. Nucleic Acids Res. **33:**2302–2309.
50. **Roostaee A, Barbar E, Lavigne P, LeHoux JG.** 2009. The mechanism of specific binding of free cholesterol by the steroidogenic acute regulatory protein: evidence for a role of the C-terminal alpha-helix in the gating of the binding site. Biosci. Rep. **29:**89–101.
51. **Poch O, Sauvaget I, Delarue M, Tordo N.** 1989. Identification of four

conserved motifs among the RNA-dependent polymerase encoding elements. EMBO J. **8:**3867–3874.
52. **Assenberg R, Delmas O, Morin B, Graham SC, De Lamballerie X, Laubert C, Coutard B, Grimes JM, Neyts J, Owens RJ, Brandt BW, Gorbalenya A, Tucker P, Stuart DI, Canard B, Bourhy H.** 2010. Genomics and structure/function studies of Rhabdoviridae proteins involved in replication and transcription. Antiviral Res. **87:**149–161.
53. **Ruigrok RW, Crepin T, Kolakofsky D.** 2011. Nucleoproteins and nucleocapsids of negative-strand RNA viruses. Curr. Opin. Microbiol. **14:**504–510.
54. **Lamb RA.** 2007. Mononegavirales, p 1357–1361. *In* Knipe DM, Howley PM, Griffin DE, Lamb RA, Martin MA, Roizman B, Straus SE (ed), Fields virology, 5th ed, vol 1. Lippincott Williams & Wilkins, Philadelphia, PA.
55. **Castagne N, Barbier A, Bernard J, Rezaei H, Huet JC, Henry C, Da Costa B, Eleouet JF.** 2004. Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain. J. Gen. Virol. **85:**1643–1653.
56. **Reid SP, Cardenas WB, Basler CF.** 2005. Homo-oligomerization facilitates the interferon-antagonist activity of the ebolavirus VP35 protein. Virology **341:**179–189.
57. **Leyrat C, Jensen MR, Ribeiro EA, Gérard F, Ruigrok R, Blackledge M, Jamin M.** 2011. The N^0-binding region of the vesicular stomatitis virus phosphoprotein is globally disordered but contains transient α-helices. Protein Sci. **20:**542–556.
58. **Leyrat C, Yabukarski F, Tarbouriech N, Ribeiro EA, Jr, Jensen MR, Blackledge M, Ruigrok RW, Jamin M.** 2011. Structure of the vesicular stomatitis virus N-P complex. PLoS Pathog. **7:**e1002248. doi:10.1371/journal.ppat.1002248.
59. **Tarbouriech N, Curran J, Ruigrok RW, Burmeister WP.** 2000. Tetrameric coiled coil domain of Sendai virus phosphoprotein. Nat. Struct. Biol. **7:**777–781.
60. **Pringle CR.** 1997. The order Mononegavirales—current status. Arch. Virol. **142:**2321–2326.
61. **Hogenhout SA, Redinbaugh MG, Ammar el-D.** 2003. Plant and animal rhabdovirus host range: a bug's view. Trends Microbiol. **11:**264–271.
62. **Green TJ, Luo M.** 2009. Structure of the vesicular stomatitis virus nucleocapsid in complex with the nucleocapsid-binding domain of the small polymerase cofactor, P. Proc. Natl. Acad. Sci. U. S. A. **106:**11721–11726.
63. **Ribeiro EA, Leyrat C, Gérard FC, Albertini AA, Falk C, Ruigrok RW, Jamin M.** 2009. Binding of rabies virus polymerase cofactor to recombinant circular nucleoprotein-RNA complexes. J. Mol. Biol. **394:**558–575.
64. **Morin B, Rahmeh AA, Whelan SP.** 2012. Mechanism of RNA synthesis initiation by the vesicular stomatitis virus polymerase. EMBO J. **31:**1320–1329.
65. **Stillman EA, Whitt M.** 1999. Transcript initiation and 5′-end modifications are separable events during vesicular stomatitis virus transcription. J. Virol. **73:**7199–7209.
66. **Plaxco KW, Simons KT, Baker D.** 1998. Contact order, transition state placement and the refolding rates of single domain proteins. J. Mol. Biol. **277:**985–994.
67. **Jensen MR, Communie G, Ribeiro EA, Jr, Martinez N, Desfosses A, Salmon L, Mollica L, Gabel F, Jamin M, Longhi S, Ruigrok RW, Blackledge M.** 2011. Intrinsic disorder in measles virus nucleocapsids. Proc. Natl. Acad. Sci. U. S. A. **108:**9839–9844.
68. **Gely S, Lowry DF, Bernard C, Jensen MR, Blackledge M, Costanzo S, Bourhis JM, Darbon H, Daughdrill G, Longhi S.** 2010. Solution structure of the C-terminal X domain of the measles virus phosphoprotein and interaction with the intrinsically disordered C-terminal domain of the nucleoprotein. J. Mol. Recognit. **23:**435–447.
69. **Jensen MR, Bernado P, Houben K, Blanchard L, Marion D, Ruigrok RW, Blackledge M.** 2010. Structural disorder within Sendai virus nucleoprotein and phosphoprotein: insight into the structural basis of molecular recognition. Protein Pept. Lett. **17:**952–960.
70. **Kingston RL, Hamel DJ, Gay LS, Dahlquist FW, Matthews BW.** 2004. Structural basis for the attachment of a paramyxoviral polymerase to its template. Proc. Natl. Acad. Sci. U. S. A. **101:**8301–8306.
71. **Galloux M, Tarus B, Blazevic I, Fix J, Duquerroy S, Eleouet JF.** 2012. Characterization of a viral phosphoprotein binding site on the surface of the respiratory syncytial nucleoprotein. J. Virol. **86:**8375–8387.
72. **Tran TL, Castagne N, Bhella D, Varela PF, Bernard J, Chilmonczyk S, Berkenkamp S, Benhamo V, Grznarova K, Grosclaude J, Nespoulos C, Rey FA, Eleouet JF.** 2007. The nine C-terminal amino acids of the respiratory syncytial virus protein P are necessary and sufficient for binding to ribonucleoprotein complexes in which six ribonucleotides are contacted per N protein protomer. J. Gen. Virol. **88:**196–206.
73. **Prins KC, Binning JM, Shabman RS, Leung DW, Amarasinghe GK, Basler CF.** 2010. Basic residues within the ebolavirus VP35 protein are required for its viral polymerase cofactor function. J. Virol. **84:**10581–10591.
74. **Takacs AM, Das T, Banerjee AK.** 1993. Mapping of interacting domains between the nucleocapsid protein and the phosphoprotein of vesicular stomatitis virus by using a two-hybrid system. Proc. Natl. Acad. Sci. U. S. A. **90:**10375–10379.
75. **Barik S, Banerjee AK.** 1992. Sequential phosphorylation of the phosphoprotein of vesicular stomatitis virus by cellular and viral protein kinases is essential for transcription activation. J. Virol. **66:**1109–1118.
76. **Das T, Pattnaik AK, Takacs AM, Li T, Hwang LN, Banerjee AK.** 1997. Basic amino acid residues at the carboxy-terminal eleven amino acid region of the phosphoprotein (P) are required for transcription but not for replication of vesicular stomatitis virus genome RNA. Virology **238:**103–114.
77. **Moseley GW, Filmer RP, DeJesus MA, Jans DA.** 2007. Nucleocytoplasmic distribution of rabies virus P-protein is regulated by phosphorylation adjacent to C-terminal nuclear import and export signals. Biochemistry **46:**12053–12061.
78. **Pasdeloup D, Poisson N, Raux H, Gaudin Y, Ruigrok RW, Blondel D.** 2005. Nucleocytoplasmic shuttling of the rabies virus P protein requires a nuclear localization signal and a CRM1-dependent nuclear export signal. Virology **334:**284–293.
79. **Blondel D, Regad T, Poisson N, Pavie B, Harper F, Pandolfi PP, De The H, Chelbi-Alix MK.** 2002. Rabies virus P and small P products interact directly with PML and reorganize PML nuclear bodies. Oncogene **21:**7957–7970.
80. **Brzozka K, Finke S, Conzelmann KK.** 2006. Inhibition of interferon signaling by rabies virus phosphoprotein P: activation-dependent binding of STAT1 and STAT2. J. Virol. **80:**2675–2683.
81. **Vidy A, Chelbi-Alix M, Blondel D.** 2005. Rabies virus P protein interacts with STAT1 and inhibits interferon signal transduction pathways. J. Virol. **79:**14411–14420.
82. **Kabsch W.** 2010. XDS. Acta Crystallogr. D Biol. Crystallogr. **66:**125–132.
83. **Heldrick GM.** 2010. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr. D Biol. Crystallogr. **66:**479–485.
84. **Terwilliger TC.** 2003. Automated main-chain model building by template matching and iterative fragment extension. Acta Crystallogr. D Biol. Crystallogr. **59:**38–44.
85. **Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA.** 2011. REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. **67:**355–367.